36-720: Graphical Models

Brian Junker

September 17, 2007

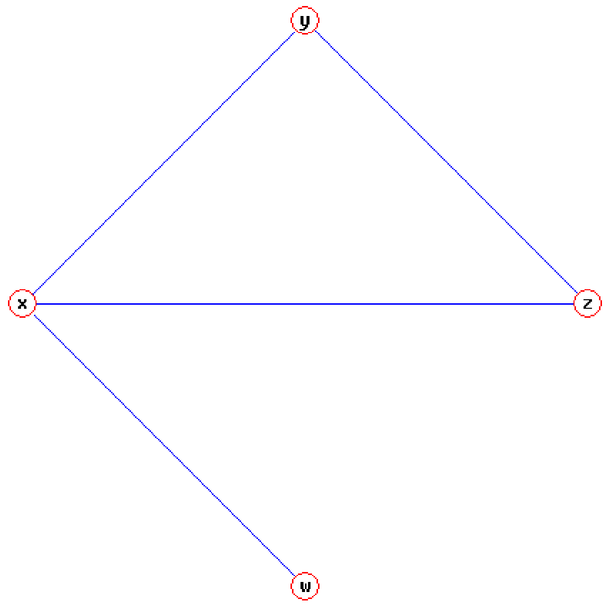
- Undirected Graphs and Conditional Independence
- Generators and Graphs
- Graphical Models
- Log-Linear Graphical Models
- Example
- Decomposabe Models

References:

- Christensen Ch 5 is a good introduction.
- Edwards, D. (2000). *Introduction to graphical modeling.* 2nd Ed. NY: Springer. (see also `http://www.hypergraph.dk` for MIM).
- Whittaker, J. (1990). *Graphical models in applied multivariate statistics.* New York: Wiley.
- Lauritzen, S. L. (1996). *Graphical models.* Oxford: Clarendon Press.

Undirected Graphs and Conditional Independence

An *undirected graph* is a set of *nodes* and *edges*:



We are interested in using such graphs to represent conditional independence in multivariate distributions:

- Nodes are random variables
- Edges represent "dependence"

The *Markov Property* provides a set of semantics for this.

The Markov Property (Axiom, really!)

- *Pairwise Markov Property:* X and Y have *no edge between them*, iff $X \perp\!\!\!\perp Y \mid$ (the rest) [where "the rest" means all other variables in the model]. All other edges are present.

- *Global Markov Property:* Two sets of variables $\mathcal{U}$ and $\mathcal{V}$ are *separated* by a third set of variables $\mathcal{W}$, if and only if $\mathcal{U} \perp\!\!\!\perp \mathcal{V} \mid \mathcal{W}$.

- *Local Markov Property:* X is conditionally independent of its non-neighbors in the graph, given its neighbors. More formally, let $\mathcal{V}$ be all variables, in the graph, and let $bd(X)$ be all the variables that are connected by one edge to X. Then $X \perp\!\!\!\perp (\mathcal{V} \setminus \{X \cup bd(X)\}) \mid bd(X)$.

- Pearl and Paz (1986): If the joint density of the variables is strictly positive wrt a product measure, then the three Markov properties are equivalent.

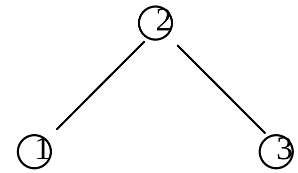
- In the graph above, we can read $W \perp\!\!\!\perp Y|(X,Z)$, $W \perp\!\!\!\perp Z|(X,Y)$, for example. Also, $W \perp\!\!\!\perp (Y,Z)|X$.

Generators and Graphs

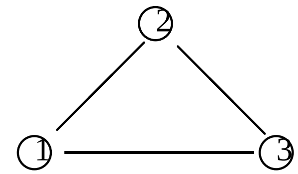
Graphs are an alternative to generators for describing dependence relations in log-linear models. Here are some examples.

Generator — *Graph*

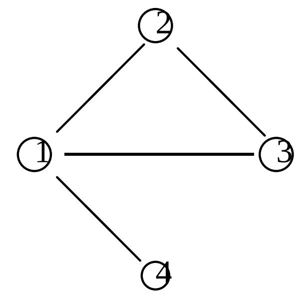
[12][23]



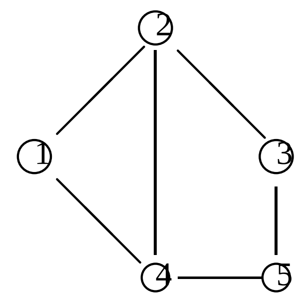
[12][13][23]



[123][14]

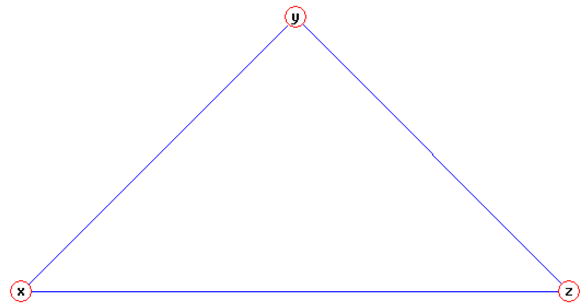


[124][23][35][45]



Graphical Models

A problem with graphs is that there is not a way to distiguish between two-way interactions and higher-w



may represent either three two-way dependencies, one three-way dependency, or both.

Definition A *clique* is a maximal complete subgraph (complete = has all possible edges; maximal = not a proper subgraph of some larger complete graph).

Definition A probability model is *graphical* if, for every clique in its conditional independence graph, all possible dependencies implied by the clique are present.

Another way to describe graphical models is that whenever all two-way dependencies are present between a set of variables, then all higher-way dependencies are also present.

Log-Linear Graphical Models

Graphical models have the nice feature that their dependence structure can be completely determined from an undirected graph and the Markov property, *and can be specified by listing the cliques in the graph.*

Typically we work within a family of probability models, like the family of *hierarchical log-linear models* or the family of *multivariate normal* (Gaussian) models, and we are interested in the graphical models within that family.

- For hierarchical log-linear models, *not all models are graphical*. For example [12][13][23] is non-graphical, while [123] and [12][13] are graphical.
- A common way to interpret a log-linear model is to *find the smallest graphical model containing it*, and interpret the (conditional) independence relationships in that graphical model.
- It turns out that *all Gaussian models are graphical*, because all higher-way interactions are determined by two-way interactions in the multivariate normal.

Example

(Handout: Edwards table, p. 23)

Decomposabe Models

One nice feature of graphical models is that they lead immediately to two important ideas in simplifying log-linear models. They are

- Decomposable models
- Collapsible models

Decomposable models are graphical models for which closed-form MLE's exist.

Collapsibility is closely related to (the non-occurrence of) Simpson's paradox like phenomena.

More on this next time, but following is a taste...

- It turns out that decomposable graphical models are exactly those models that are "triangulated":
  - A *chordless cycle* is a cycle or closed loop with no shortcuts.
  - A *triangulated graph* is one who chordless cycles contain no more than three vertices.
- For a triangulated graphical model it is easy to read off the MLE's:
  - The cell counts will have the form

$$m_{ijklm\cdots} = \frac{\textit{minimal sufficient margins}}{\textit{separator margins}}$$

    where
    - The *"minimal sufficient margins"* are the margins with fixed indices corresponding to the terms in the generator (cliques in the graph!); and
    - the *"separator margins"* are the margins with fixed indices corresponding to variables common to terms in the generator (these terms are maximal sets of vertices separating the graph into disconnected parts).

Examples in class.